

# ISO/IEC 8859-1 - Wikipedia

W en.wikipedia.org/wiki/ISO/IEC\_8859-1

For the Unicode block also called "Latin 1", see [Latin-1 Supplement \(Unicode block\)](#). For the 1990s-era misnomer "ISO-8859-1", see [Windows-1252](#).



ISO/IEC 8859-1 code page layout

MIME / IANA	ISO-8859-1
Alias(es)	iso-ir-100, csISOLatin1, latin1, l1, IBM819, CP819
Language(s)	<a href="#">English</a> , <a href="#">various others</a>
Standard	<a href="#">ISO/IEC 8859</a>
Classification	<a href="#">Extended ASCII</a> , <a href="#">ISO/IEC 8859</a>
Extends	<a href="#">US-ASCII</a>
Based on	<a href="#">DEC MCS</a>
Succeeded by	<ul style="list-style-type: none"><li><a href="#">UTF-8</a></li><li><a href="#">UTF-16</a></li></ul>
Other related encoding(s)	<ul style="list-style-type: none"><li><a href="#">ISO/IEC 8859-15</a></li><li><a href="#">Windows-1252</a></li><li><a href="#">BraSCII</a></li></ul>

ISO/IEC 8859-1:1998

**ISO/IEC 8859-1:1998**, *Information technology—8-bit single-byte coded graphic character sets—Part 1: Latin alphabet No. 1*, is part of the [ISO/IEC 8859](#) series of [ASCII](#)-based standard [character encodings](#), first edition published in 1987. ISO/IEC 8859-1 encodes what it refers to as "**Latin alphabet no. 1**", consisting of 191 [characters](#) from the [Latin script](#). This character-encoding scheme is used throughout the [Americas](#), [Western Europe](#), [Oceania](#), and much of [Africa](#). It is the basis for some popular 8-bit character sets and the first two blocks of characters in [Unicode](#).

As of July 2024, 1.2% of all web sites use ISO/IEC 8859-1.<sup>[1][2]</sup> It is the most declared single-byte character encoding, but as Web browsers and the HTML5 standard<sup>[3]</sup> interpret them as the superset Windows-1252, these documents may include characters from that set. Depending on the country or language, website use can be higher than the global average, in Brazil it is at 3.4%,<sup>[4]</sup> and in Germany at 2.7%.<sup>[5][6]</sup>

ISO-8859-1 was (according to the standard, at least) the default encoding of documents delivered via HTTP with a MIME type beginning with `text/`, the default encoding of the values of certain descriptive HTTP headers, and defined the repertoire of characters allowed in HTML 3.2 documents. It is specified by many other standards. In practice, the superset encoding Windows-1252 is the more likely effective default<sup>[7]</sup> and it is increasingly common for standards to (at least unofficially) default to UTF-8.

**ISO-8859-1** is the IANA preferred name for this standard when supplemented with the C0 and C1 control codes from ISO/IEC 6429. The following other aliases are registered: **iso-ir-100**, **csISOLatin1**, **latin1**, **I1**, **IBM819**, **Code page 28591** a.k.a. **Windows-28591** is used for it in Windows.<sup>[8]</sup> IBM calls it **code page 819** or **CP819** (**CCSID 819**).<sup>[9][10][11][12]</sup> Oracle calls it **WE8ISO8859P1**.<sup>[13]</sup>

## Coverage

---

See also: Latin-script alphabet

Each character is encoded as a single eight-bit code value. These code values can be used in almost any data interchange system to communicate in the following languages (while it may exclude correct quotation marks such as for many languages including German and Icelandic):

## Modern languages with complete coverage

---

- Afrikaans
- Albanian
- Basque
- Breton
- Corsican
- English
- Faroese
- Galician
- Icelandic
- Ido
- Irish
- Indonesian
- Italian
- Leonese
- Lojban
- Luxembourgish<sup>[a]</sup>
- Malay<sup>[b]</sup>

- [Manx](#)
- [Norwegian](#)<sup>[c]</sup>
- [Occitan](#)
- [Portuguese](#)<sup>[d]</sup>
- [Rhaeto-Romanic](#)
- [Rotokas](#)
- [Scottish Gaelic](#)
- [Scots](#)
- [Southern Sami](#)
- [Spanish](#)
- [Swahili](#)
- [Swedish](#)
- [Tagalog](#)
- [Toki Pona](#)
- [Walloon](#)

## Notes

1. <sup>^</sup> Basic classical orthography
2. <sup>^</sup> [Rumi script](#)
3. <sup>^</sup> Bokmål and Nynorsk
4. <sup>^</sup> European and Brazilian

## Languages with incomplete coverage

ISO-8859-1 was commonly used for certain languages, even though it lacks characters used by these languages. In most cases, only a few letters are missing or they are rarely used, and they can be replaced with characters that are in ISO-8859-1 using some form of typographic approximation. The following table lists such languages.

Language	Missing characters	Typical workaround	Supported by
<a href="#">Catalan</a>	Ł, ł (deprecated)	L·, l·	
<a href="#">Danish</a>	Ø, ø (the accent is optional and ø is very rare)	Ø, ø or øe	
<a href="#">Dutch</a>	Ĳ, ĳ (debatable); ĵ in emphasized words like "blĳf"	<u>digraphs</u> IJ, ij or ŷ; blĳf	
<a href="#">Estonian</a> , <a href="#">Finnish</a>	Š, š, Ž, ž (only present in loanwords)	Sh, sh, Zh, zh	<a href="#">ISO-8859-15</a> , <a href="#">Windows-1252</a>

<u>French</u>	<u>Œ</u> , <u>œ</u> , and the very rare <u>Ÿ</u>	<u>digraphs</u> OE, oe; Y or Ý	<u>ISO-8859-15</u> , <u>Windows-1252</u>
<u>German</u>	<u>ß</u> (capital ß, used only in all capitals)	<u>digraph</u> SS or SZ	
<u>Hungarian</u>	<u>Ő</u> , <u>ő</u> , <u>Ű</u> , <u>ű</u>	<u>Ö</u> , <u>ö</u> , <u>Ü</u> , <u>ü</u> <u>Ő</u> , <u>ő</u> , <u>Ű</u> , <u>ű</u> (the characters replaced in <u>8859-2</u> )	<u>ISO-8859-2</u> , <u>Windows-1250</u>
<u>Irish</u> ( <u>traditional orthography</u> )	<u>Ḃ</u> , <u>ḃ</u> , <u>Ḇ</u> , <u>ḇ</u> , <u>Ḋ</u> , <u>ḋ</u> , <u>Ḟ</u> , <u>ḟ</u> , <u>Ḡ</u> , <u>ḡ</u> , <u>Ḣ</u> , <u>ḣ</u> , <u>Ḗ</u> , <u>ḗ</u> , <u>Ḣ</u> , <u>ḣ</u> , <u>Ṳ</u> , <u>ṳ</u>	Bh, bh, Ch, ch, Dh, dh, Fh, fh, Gh, gh, Mh, mh, Ph, ph, Sh, sh, Th, th	<u>ISO-8859-14</u>
<u>Welsh</u>	<u>Ŵ</u> , <u>ŵ</u> , <u>Ŷ</u> , <u>ŷ</u> , <u>Ŷ</u> , <u>ŷ</u> , <u>Ŵ</u> , <u>ŵ</u> , <u>Ŷ</u> , <u>ŷ</u> , <u>Ŷ</u> , <u>ŷ</u>	W, w, Y, y, Ý, ý	<u>ISO-8859-14</u>

The letter ÿ, which appears in French only very rarely, mainly in city names such as L'Haÿ-les-Roses and never at the beginning of words, is included only in lowercase form. The slot corresponding to its uppercase form is occupied by the lowercase letter ß from the German language, which did not have an uppercase form at the time when the standard was created.

## Quotation marks

For some languages listed above, the correct typographical quotation marks are missing, as only « », " ", and ' ' are included. Also, this scheme does not provide for oriented (6- or 9-shaped) single or double quotation marks. Some fonts will display the spacing grave accent (0x60) and the apostrophe (0x27) as a matching pair of oriented single quotation marks (see Quotation mark § Typewriters and early computers), but this is not considered part of the modern standard.

## History

ISO 8859-1 was based on the Multinational Character Set (MCS) used by Digital Equipment Corporation (DEC) in the popular VT220 terminal in 1983. It was developed within the European Computer Manufacturers Association (ECMA), and published in March 1985 as ECMA-94,<sup>[14]</sup> by which name it is still sometimes known. The second edition of ECMA-94 (June 1986)<sup>[15]</sup> also included ISO 8859-2, ISO 8859-3, and ISO 8859-4 as part of the specification.

The original draft of ISO 8859-1 placed French Œ and œ at code points 215 (0xD7) and 247 (0xF7), as in the MCS. However, the delegate from France, being neither a linguist nor a typographer, falsely stated that these are not independent French letters on their own, but mere ligatures (like *fi* or *fl*), supported by the delegate team from Bull Publishing Company, who regularly did not print French with Œ/œ in their house style at the time. An anglophone delegate from Canada insisted on retaining Œ/œ but was rebuffed by the

French delegate and the team from Bull. These code points were soon filled with × and ÷ under the suggestion of the German delegation. Support for French was further reduced when it was again falsely stated that the letter ÿ is "not French", resulting in the absence of the capital Ÿ. In fact, the letter ÿ is found in a number of French proper names, and the capital letter has been used in dictionaries and encyclopedias.<sup>[16]</sup> These characters were added to ISO/IEC 8859-15:1999. BraSCII matches the original draft.

In 1985, Commodore adopted ECMA-94 for its new AmigaOS operating system.<sup>[17]</sup> The Seikosha MP-1300AI impact dot-matrix printer, used with the Amiga 1000, included this encoding.

In 1990, the first version of Unicode used the code points of ISO-8859-1 as the first 256 Unicode code points.

In 1992, the IANA registered the character map **ISO\_8859-1:1987**, more commonly known by its preferred MIME name of **ISO-8859-1** (note the extra hyphen over ISO 8859-1), a superset of ISO 8859-1, for use on the Internet. This map assigns the C0 and C1 control codes to the unassigned code values thus provides for 256 characters via every possible 8-bit value.

## Code page layout

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0x																
1x																
2x	<u>SP</u>	!	"	#	\$	%	&	'	(	)	*	±	,	=	~	/
3x	<u>Q</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>:</u>	<u>;</u>	<u>≤</u>	<u>≡</u>	<u>≥</u>	<u>?</u>
4x	<u>@</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>H</u>	<u>I</u>	<u>J</u>	<u>K</u>	<u>L</u>	<u>M</u>	<u>N</u>	<u>O</u>
5x	<u>P</u>	<u>Q</u>	<u>R</u>	<u>S</u>	<u>T</u>	<u>U</u>	<u>V</u>	<u>W</u>	<u>X</u>	<u>Y</u>	<u>Z</u>	[	\	]	<u>^</u>	<u>_</u>
6x	<u>`</u>	<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	<u>f</u>	<u>g</u>	<u>h</u>	<u>i</u>	<u>j</u>	<u>k</u>	<u>l</u>	<u>m</u>	<u>n</u>	<u>o</u>
7x	<u>p</u>	<u>q</u>	<u>r</u>	<u>s</u>	<u>t</u>	<u>u</u>	<u>v</u>	<u>w</u>	<u>x</u>	<u>y</u>	<u>z</u>	{		}	<u>~</u>	
8x																
9x																
Ax	<u>NBSP</u>	<u>ı</u>	<u>ø</u>	<u>£</u>	<u>¤</u>	<u>¥</u>	<u>ı</u>	<u>§</u>	<u>¨</u>	<u>©</u>	<u>ª</u>	<u>«</u>	<u>»</u>	<u>SHY</u>	<u>®</u>	<u>™</u>
Bx	<u>°</u>	<u>±</u>	<u>²</u>	<u>³</u>	<u>´</u>	<u>µ</u>	<u>¶</u>	<u>:</u>	<u>;</u>	<u>¹</u>	<u>º</u>	<u>»</u>	<u>¼</u>	<u>½</u>	<u>¾</u>	<u>¿</u>
Cx	<u>À</u>	<u>Á</u>	<u>Â</u>	<u>Ã</u>	<u>Ä</u>	<u>Å</u>	<u>Æ</u>	<u>Ç</u>	<u>È</u>	<u>É</u>	<u>Ê</u>	<u>Ë</u>	<u>Ì</u>	<u>Í</u>	<u>Î</u>	<u>Ï</u>
Dx	<u>Ð</u>	<u>Ñ</u>	<u>Ò</u>	<u>Ó</u>	<u>Ô</u>	<u>Õ</u>	<u>Ö</u>	<u>×</u>	<u>Ø</u>	<u>Ù</u>	<u>Ú</u>	<u>Û</u>	<u>Ü</u>	<u>Ý</u>	<u>Þ</u>	<u>ß</u>

Ex	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
Fx	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Undefined

Symbols and punctuation

Undefined in the first release of ECMA-94 (1985).<sup>[14]</sup> In the original draft Æ was at 0xD7 and œ was at 0xF7.

## ISO/IEC 8859-1

## Similar character sets

Main article: [Western Latin character sets \(computing\)](#)

## ISO/IEC 8859-15

[ISO/IEC 8859-15](#) was developed in 1999, as an update of ISO/IEC 8859-1. It provides some characters for French and Finnish text and the [euro sign](#), which are missing from ISO/IEC 8859-1. This required the removal of some infrequently used characters from ISO/IEC 8859-1, including fraction symbols and letter-free diacritics: ◻, ı, ˆ, ˘, ˙, ¼, ½, and ¾. Ironically, three of the newly added characters (€, æ, and ÿ) had already been present in [DEC's 1983 Multinational Character Set](#) (MCS), the predecessor to ISO/IEC 8859-1 (1987). Since their original code points were now reused for other purposes, the characters had to be reintroduced under different, less logical code points.

ISO-IR-204, a more minor modification (called **code page 61235** by FreeDOS),<sup>[18]</sup> had been registered in 1998, altering ISO-8859-1 by replacing the [universal currency sign](#) (¤) with the euro sign<sup>[19]</sup> (the same substitution made by ISO-8859-15).

## Windows-1252

The popular [Windows-1252](#) character set adds all the missing characters provided by [ISO/IEC 8859-15](#), plus a number of typographic symbols, by replacing the rarely used C1 controls in the range 128 to 159 ([hex](#) 80 to 9F). It is very common to mislabel Windows-1252 text as being in ISO-8859-1. A common result was that all the quotes and apostrophes (produced by "smart quotes" in word-processing software) were replaced with question marks or boxes on non-Windows operating systems, making text difficult to read. Many Web browsers and e-mail clients will interpret ISO-8859-1 control codes as Windows-1252 characters, and that behavior was later standardized in [HTML5](#).<sup>[20]</sup>

## Mac Roman

The [Apple Macintosh](#) computer introduced a character encoding called [Mac Roman](#) in 1984. It was meant to be suitable for Western European [desktop publishing](#). It is a superset of ASCII, and has most of the characters that are in ISO-8859-1 and all the extra

characters from Windows-1252, but in a totally different arrangement. The few printable characters that are in ISO/IEC 8859-1, but not in this set, are often a source of trouble when editing text on Web sites using older Macintosh browsers, including the last version of [Internet Explorer for Mac](#).

## Other

---

DOS has [code page 850](#), which has all printable characters that ISO-8859-1 has, albeit in a totally different arrangement, plus the most widely used [graphic characters](#) from [code page 437](#).

Between 1989<sup>[21]</sup> and 2015, [Hewlett-Packard](#) used another superset of ISO-8859-1 on many of their calculators. [This proprietary character set](#) was sometimes referred to simply as "ECMA-94" as well.<sup>[21]</sup> HP also has [code page 1053](#), which adds the medium shade (☐, U+2592) at 0x7F.<sup>[22]</sup>

Several [EBCDIC](#) code pages were purposely designed to have the same set of characters as ISO-8859-1, to allow easy conversion between them.

## See also

---

## References

---

1. <sup>^</sup> [Cowan, John; Soltano, Sam \(August 2014\). "Source of character encoding statistics?". W3Techs. Archived from the original on 4 April 2024.](#)
2. <sup>^</sup> ["Encoding". WHATWG. 27 January 2015. sec. 5.2 Names and labels. Archived from the original on 4 February 2015. Retrieved 4 February 2015.](#)
3. <sup>^</sup> ["Code page 819 information document". Archived from the original on 2017-01-16.](#)
4. <sup>^</sup> ["CCSID 819 information document". Archived from the original on 2016-03-27.](#)
5. <sup>^</sup> [Baird, Cathy; Chiba, Dan; Chu, Winson; Fan, Jessica; Ho, Claire; Law, Simon; Lee, Geoff; Linsley, Peter; Matsuda, Keni; Oscroft, Tamzin; Takeda, Shige; Tanaka, Linus; Tozawa, Makoto; Trute, Barry; Tsujimoto, Mayumi; Wu, Ying; Yau, Michael; Yu, Tim; Wang, Chao; Wong, Simon; Zhang, Weiran; Zheng, Lei; Zhu, Yan; Moore, Valarie \(2002\) \[1996\]. "Appendix A: Locale Data". \*Oracle9i Database Globalization Support Guide\* \(PDF\) \(Release 2 \(9.2\) ed.\). Oracle Corporation. Oracle A96529-01. Archived \(PDF\) from the original on 2017-02-14. Retrieved 2017-02-14.](#)

6. <sup>^</sup> <sup>a</sup> <sup>b</sup> Standard ECMA-94: 8-bit Single-Byte Coded Graphic Character Set (PDF) (1 ed.). European Computer Manufacturers Association (ECMA). March 1985 [1984-12-14]. Archived (PDF) from the original on 2016-12-02. Retrieved 2016-12-01. "[...] Since 1982 the urgency of the need for an 8-bit single-byte coded character set was recognized in ECMA as well as in ANSI/X3L2 and numerous working papers were exchanged between the two groups. In February 1984 ECMA TC1 submitted to ISO/TC97/SC2 a proposal for such a coded character set. At its meeting of April 1984 SC decided to submit to TC97 a proposal for a new item of work for this topic. Technical discussions during and after this meeting led TC1 to adopt the coding scheme proposed by X3L2. Part 1 of Draft International Standard DTS 8859 is based on this joint ANSI/ECMA proposal. [...] Adopted as an ECMA Standard by the General Assembly of Dec. 13–14, 1984. [...]"
7. <sup>^</sup> André, Jacques (1996). "ISO Latin-1, norme de codage des caractères européens? Trois caractères français en sont absents!" (PDF). *Cahiers GUTenberg* (in French) (25): 65–77. doi:10.5802/cg.205.
8. <sup>^</sup> Malyshev, Michael (2003-01-10). "Registration of new charset [Amiga-1251]". ATO-RU (Amiga Translation Organization - Russian Department). Archived from the original on 2016-12-05. Retrieved 2016-12-05.
9. <sup>^</sup> ITS Information Technology Standardization (1998-09-16). Supplementary set for Latin-1 alternative with EURO SIGN (PDF). ITSCJ/IPSJ. ISO-IR-204.
10. <sup>^</sup> van Kesteren, Anne (27 January 2015). "5.2 Names and labels". *Encoding Standard. WHATWG*. Archived from the original on 4 February 2015. Retrieved 4 February 2015.
11. <sup>^</sup> <sup>a</sup> <sup>b</sup> HP 82240B Infrared Printer (1 ed.). Corvallis, OR, USA: Hewlett-Packard. August 1989. HP reorder number 82240-90014.
12. <sup>^</sup> "Code Page 1053" (PDF). Archived from the original (PDF) on 2013-01-21.

## External links

---

- ISO/IEC 8859-1:1998
- ISO/IEC FDIS 8859-1:1998 Archived 2020-09-30 at the Wayback Machine — 8-bit single-byte coded graphic character sets, Part 1: Latin alphabet No. 1 (*draft dated February 12, 1998, published April 15, 1998*)
- Standard ECMA-94: 8-Bit Single Byte Coded Graphic Character Sets — Latin Alphabets No. 1 to No. 4 2nd edition (June 1986)
- ISO-IR 100 Right-Hand Part of Latin Alphabet No.1 (*February 1, 1986*)
- The Letter Database
- Czyborra, Roman (1998-12-01). "The ISO 8859 Alphabet Soup". Archived from the original on 2016-12-01. Retrieved 2016-12-01. [1] [2]

## Character encodings

---



---

## Early telecommunications

- Telegraph code
  - Needle
  - Morse
    - Non-Latin
    - Wabun/Kana
    - Chinese
    - Cyrillic
  - Baudot and Murray
- Fieldata
- ASCII
  - ISO/IEC 646
- BCDIC
- Teletex and Videotex/Teletext
  - T.51/ISO/IEC 6937
  - ITU T.61
  - ITU T.101
  - World System Teletext
    - background
    - sets
- Transcode

---

## ISO/IEC 8859

- Approved parts
    - -1 (Western Europe)
    - -2 (Central Europe)
    - -3 (Maltese/Esperanto)
    - -4 (North Europe)
    - -5 (Cyrillic)
    - -6 (Arabic)
    - -7 (Greek)
    - -8 (Hebrew)
    - -9 (Turkish)
    - -10 (Nordic)
    - -11 (Thai)
    - -13 (Baltic)
    - -14 (Celtic)
    - -15 (New Western Europe)
    - -16 (Romanian)
  - Abandoned parts
    - 12 (Devanagari)
  - Proposed but not approved
    - KOI-8 Cyrillic
    - Sámi
  - Adaptations
    - Welsh
    - Barents Cyrillic
    - Estonian
    - Ukrainian Cyrillic
-

---

**Bibliographic use**

- MARC-8
  - ANSEL
  - CCCI/EACC
- ISO 5426
- 5426-2
- 5427
- 5428
- 6438
- 6862

---

**National standards**

- ArmSCII
- Big5
- BraSCII
- CNS 11643
- DIN 66003
- ELOT 927
- GOST 10859
- GB 2312
- GB 12345
- GB 12052
- GB 18030
- HKSCS
- ISCII
- JIS X 0201
- JIS X 0208
- JIS X 0212
- JIS X 0213
- KOI-7
- KPS 9566
- KS X 1001
- KS X 1002
- LST 1564
- LST 1590-4
- PASCI
- Shift JIS
- SI 960
- TIS-620
- TSCII
- VISCII
- VSCII
- YUSCII

---

**ISO/IEC 2022**

- ISO/IEC 8859
  - ISO/IEC 10367
  - Extended Unix Code / EUC
-

---

**Mac OS Code pages**  
**("scripts")**

- Armenian
- Arabic
- Barents Cyrillic
- Celtic
- Central European
- Croatian
- Cyrillic
- Devanagari
- Farsi (Persian)
- Font X (Kermit)
- Gaelic
- Georgian
- Greek
- Gujarati
- Gurmukhi
- Hebrew
- Iceland
- Inuit
- Keyboard
- Latin (Kermit)
- Maltese/Esperanto
- Ogham
- Roman
- Romanian
- Sámi
- Turkish
- Turkic Cyrillic
- Ukrainian
- VT100

---

**DOS code pages**

- 437
- 668
- 708
- 720
- 737
- 770
- 773
- 775
- 776
- 777
- 778
- 850
- 851
- 852
- 853
- 855
- 856
- 857
- 858
- 859
- 860
- 861
- 862
- 863

- 864
  - 865
  - 866
  - 867
  - 868
  - 869
  - 897
  - 899
  - 903
  - 904
  - 932
  - 936
  - 942
  - 949
  - 950
  - 951
  - 1034
  - 1040
  - 1042
  - 1043
  - 1044
  - 1098
  - 1115
  - 1116
  - 1117
  - 1118
  - 1127
  - 3846
  - ABICOMP
  - CS Indic
  - CSX Indic
  - CSX+ Indic
  - CWI-2
  - Iran System
  - Kamenický.
  - Mazovia
  - MIK
-

---

**IBM AIX code pages**

- [895](#)
- [896](#)
- [912](#)
- [915](#)
- [921](#)
- [922](#)
- [1006](#)
- [1008](#)
- [1009](#)
- [1010](#)
- [1012](#)
- [1013](#)
- [1014](#)
- [1015](#)
- [1016](#)
- [1017](#)
- [1018](#)
- [1019](#)
- [1046](#)
- [1124](#)
- [1133](#)

---

**Windows code pages**

- [CER-GS](#)
- [932](#)
- [936 \(GBK\)](#)
- [950](#)
- [1169](#)
- [Extended Latin-8](#)
- [1250](#)
- [1251](#)
- [1252](#)
- [1253](#)
- [1254](#)
- [1255](#)
- [1256](#)
- [1257](#)
- [1258](#)
- [1270](#)
- [Cyrillic + Finnish](#)
- [Cyrillic + French](#)
- [Cyrillic + German](#)
- [Polytonic Greek](#)

---

**EBCDIC code pages**

- [Japanese language in EBCDIC](#)
  - [DKOI](#)
-

---

## DEC terminals (VTx)

- Multinational (MCS)
  - National Replacement (NRCS)
    - French Canadian
    - Swiss
    - Spanish
    - United Kingdom
    - Dutch
    - Finnish
    - French
    - Norwegian and Danish
    - Swedish
    - Norwegian and Danish (alternative)
  - 8-bit Greek
  - 8-bit Turkish
  - SI 960
  - Hebrew
  - Special Graphics
  - Technical (TCS)
-

---

## Platform specific

- [1052](#)
- [1053](#)
- [1054](#)
- [1055](#)
- [1056](#)
- [1057](#)
- [1058](#)
- [Acorn RISC OS](#)
- [Amstrad CPC](#)
- [Apple II](#)
- [ATASCII](#)
- [Atari ST](#)
- [BICS](#)
- [Casio calculators](#)
- [CDC](#)
- [Compucolor 8001](#)
- [Compucolor II](#)
- [CP/M+](#)
- [DEC RADIX 50](#)
- [DEC MCS/NRCS](#)
- [DG International](#)
- [Galaksija](#)
- [GEM](#)
- [GSM 03.38](#)
- [HP Roman](#)
- [HP FOCAL](#)
- [HP RPL](#)
- [SQUOZE](#)
- [LICS](#)
- [LMBCS](#)
- [MSX](#)
- [NEC APC](#)
- [NeXT](#)
- [PETSCII](#)
- [PostScript Standard](#)
- [PostScript Latin 1](#)
- [SAM Coupé](#)
- [Sega SC-3000](#)
- [Sharp calculators](#)
- [Sharp MZ](#)
- [Sinclair QL](#)
- [Teletext](#)
- [TI calculators](#)
- [TRS-80](#)
- [Ventura International](#)
- [WISCI](#)
- [XCCS](#)
- [ZX80](#)
- [ZX81](#)
- [ZX Spectrum](#)

---

## **Unicode / ISO/IEC 10646**

- [UTF-1](#)
- [UTF-7](#)
- [UTF-8](#)
- [UTF-16](#)
- [UTF-32](#)
- [UTF-EBCDIC](#)
- [GB 18030](#)
- [DIN 91379](#)
- [BOCU-1](#)
- [CESU-8](#)
- [SCSU](#)
- [TACE16](#)
- [Comparison of Unicode encodings](#)

---

## **TeX typesetting system**

- [Cork](#)
- [LY1](#)
- [OML](#)
- [OMS](#)
- [OT1](#)

---

## **Miscellaneous code pages**

- [ABICOMP](#)
- [ASMO 449](#)
- [Digital encoding of APL symbols](#)  
[ISO-IR-68](#)
- [ARIB STD-B24](#)
- [Fieldata](#)
- [HZ](#)
- [IEC-P27-1](#)
- [INIS](#)
  - [7-bit](#)
  - [8-bit](#)
- [ISO-IR-169](#)
- [ISO 2033](#)
- [KOI](#)
  - [KOI8-R](#)
  - [KOI8-RU](#)
  - [KOI8-U](#)
- [Mojikyo](#)
- [SEASCII](#)
- [Stanford/ITS](#)
- [Symbol](#)
- [TRON](#)
- [Unified Hangul Code](#)

---

## **Control character**

- [Morse prosigns](#)
  - [C0 and C1 control codes](#)
    - [ISO/IEC 6429](#)
    - [JIS X 0211](#)
  - [Unicode control, format and separator characters](#)
  - [Whitespace characters](#)
-



---

**Related topics**

- [CCSID](#)
- [Character encodings in HTML](#)
- [Charset detection](#)
- [Han unification](#)
- [Hardware code page](#)
- [MICR code](#)
- [Mojibake](#)
- [Variable-length encoding](#)

---

 [Character sets](#)